

FIG. 1

tgagaccaagccaacatctatgcagacttgtcagcagccgaatgtcatcctggcaggcggtccctgggacagtgcagtgtca
cttggcagggataccagactaagagcagtgaaatgcattggacttatgtcagtggtagatgacaatgactgtaatgca
gcaactagaccaactgatacccaggactgtgaattaccatcatgtcatcctccccagctgccccgaaacgaggagaagcacata
cagtgcaccaagaacccagtggcattggcttgggacccatgtcagccacttgtggaaaaggtagccggatgagatacgtca
5 5 gctgccgagatgagaatggcttgtggctgacgagagtgcctgtctaccctgcctagaccagtggcaaaggaaatgttctgt
acaccctgtggcaatggaaaggccttggactggagctttgtgaccctgtggcaaggtagggcaacccggcaagtgt
tgtcaactacagtgaccacgtgatcgatcgagttgtgaccaggattatcccagaaactgaccaggactgttccatgtcac
catgcctcaaaggacccagacagtggcttagctcagcacccctccaaaatgaggactatgtccccggagcgcagccagcc
10 10 cgcacccatgtcggtggaaaccagtggagaactggccctggggagcatgttccagttacactgtgtctggcgatcccagcc
tgggttgtatgtcaggatgaaaatggatacaccgcacactgtgtggagagaataaaacctgtgagcaagagccctgt
ccggcccttgcctcagttggcttatggcaactggggagagtgcactaagctgtgtggaggcataagaacaagactgg
tgtcagcggtccaacggtaacgggtttccagatttgcatt
tgcttgcacacacgcgtcatggagttactggcccttggagctgtgtctgtcttttttttttttttttttttttttttt
tttactgcatt
15 15 ggaggaagatgccccaaatggaaagctggcgcttggagtcagtgtctgtgtccctgtggccgaggctacagc
ctgtcagatcggaacacacaaaatagccagagagaccgagtgcaacccatcacccagccggactcggaacgc
cactgggttcccttctacacttggaggcagaggaatggcaagaatgcaccaagacccctccaggttcc
gtgtgtgtgatgacaacaaaaacggatgtcatggggcactggggcactggggcactggggcactgggg
gcaaccctgcgagttatgtctggatcacaggagaatggtcagaggatccctggactgt
20 20 ctgagagtggcagaggatgagtggaggatgagtggcaggatgtggagactgtggctaccgc
gtgtgtttctgacaagtccctcagttcccaagctagaattccctgtatgcaaaagcggg
cccttcaggctacattctgtgattcccttgcatttgcctattggcataaagaagaattt
gctgtgaggtgcctgaagacagttaaatattgttagtgc
25 25 tttcaagagaggctaaaaatctggacttagtgc
gtaaggggcaaaacaaaacccttcatgaaccagatgtgtgc
aatatccgaagtatctcaagtac

FIG. 2

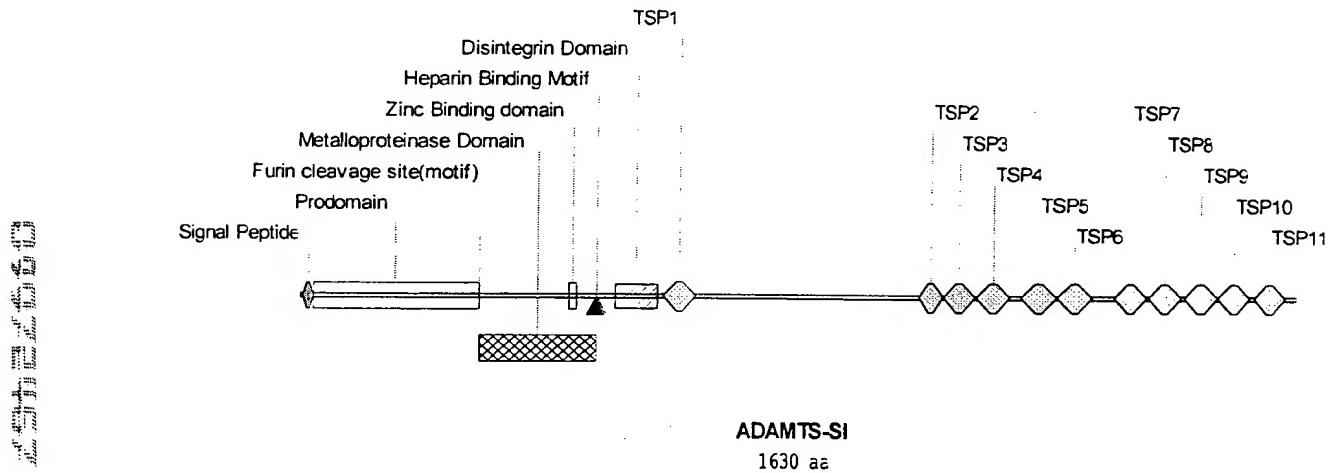
5 mqfvswatlltlvrslaemgspdaaaavrkdrlhprqvklletlseyeivspirvnalgepfptnvhfkrrsins
atdpwpafasssssstssqahyrlsafgqqflfnltanagfiapltvtllgtgvnqtkfyseeeaelkhcfykgyvntnsehta
vislcsgmgtfrshdgdyfieplqsmdeqedeeeqnkphiyyrsapqrepstgrhacdtehknrhsdkkktrarkwgerinl
agdvaalnsglateafsaygnktndtrekrthrrtckrflysprfvevlvvadnrmvsyhgenlqhyiltlmsivasiykdpsignal
inivivnlivihneqdgpsisfnaqttlknfcqwqhsknspggihdtavlltrqdicrahdkcdtlglaelgticdpyscsise
dsglstaftiahelghvfmphddnnkceegvkspqhvmaptnfytnpwmwsksrkyitefldtgygecllnepesrpypplv
10 qlpgilynvnkqcelifgpgsqvcpcymmqrrlwcnvngvhkcrtqhtpwadgtecepgkhckygfccvpkemdvpvtwgswgw
spfgtcsrtcggiktairecnrpepkngkycvgrmmkfksntepc1kqkrdrdeqcahfdgkhfnningllpnrvwpkysgi
lmkdrcklfcrvagnattyqlrdrvidgtcpqdndicvqqlcrqagcdhvlnskarrdkcgvcggdnsscktvagtfntvhgy
ntvvripagatnidvrqhsfsgetdddnylalsssksgefflngnfvvtmakreirignavveysgsetaverinstdrieqelllq
vlsvgklynnpdvrysfnipiedkpqqfywnshgpwqacsckpcqgerkrklvctresdqltvsdqrcdrlpqpghitepcgtdcdlr
15 whvasrsecaqcglyrtldiycakeysrldgktekvddgcscshpkpnsrekcsgecntggwrysawtecskscdggqriraic
vntrndvlddskcthqekvtigrcsefpccpqwksgdwseclvtcgkghkhrqvwqfcqfgedrlndrmcdpetkptsmqtcqqpecas
wqagpwgqcsvtcqgyqlravkciigtymsvvddndcnaatrptdtqdcelpschpppaapetrrrstysaprtqwrfgswtpcsa
tcgkgrmryvscrdengsvadesacatlprpvakeecsvtpcgcqwkaldwsscsvtcgqgratrqvmcvnydhvidrsecdqdy
ipetdqdcsmspcpqrtpdsglaqhpfqnedyrprsasprsrthvlggnqwrtpwgacsstcaggsqrwwvcqdengytandcve
20 rikpdeqracesgpcpqwaygnwgecklcgggirtrlvvcqrsngerfpdlsceildkppdreqcnthacphdaawstgpwsscs
vscgrghkqrnvycmaqdshlesdyckhlakphghrkcrgrcpkwkagawsqcsvscgrgvqqrhvgcqigthkiaretecnpv
trpeserdcqgprcplytwraeewqectktcgegsryrkvvcvddnknevvhgarcdvskrpdrescslqceyvwitgewsevps
wel*

25

Fig. 3

Domain structure of ADAMTS-SI

Signal peptide (1-18), Prodomain (19-287), Furin cleavage site (288), Metalloproteinase domain (289-478), Zinc binding motif (434-446), Heparin binding motif (478-482), Disintegrin domain (509-578), Thrombospondin motif (589-642), Spacer region (643-1014), Thrombospondin submotifs (1001-1053, 1056-1108, 1111-1165, 1186-1239, 1240-1295, 1332-1383, 1386-1439, 1445-1500, 1501-1554, 1559-1612)



B

1 CCGGAATTCC CGGGTCGACC CACGCGTCCG GCCCCCATT CAAGAAGCCC CTCAGCTATC
GGCCTTAAGG GCCCAGCTGG GTGCGCAGGC CGGGGGTAA TTCTTCGGC GAGTCGATAG

61 CCGGCCAGCA CAGGGCGCCC GGCGCGCTC GGAGCGCAAG TTCTCGCCT TCTCTGCC
GGCGGTCTG GTCCCGCGG CGCGCGGAG CCTCGCGTTC AAGGAGCGGA AGAGGACGGG

121 GCTCGCTGGG CATTATGCGG CCAAGCAGCC GAGCCCCAGT CCTCTCCTC CTCTGCTCC
CGAGCGACCC GTAATACGCC GGTCGTCGG CTGGGGTCA GGAGGAGGAG GAGGACGAGG

181 TCCGGCTCCT CCTCGGGCCC GAGCGGCTCA GCTCTCGCA GGCGGGGGCG TTGCTCAGCC
AGGGCGAGGA GGACGCCGGG CTCGCCAGT CGAGAGCCGT CCAGCGCCGC AACGAGTCGG

241 GAGCGCAGAC GGGACCTCG CAGCGAGACC TCAGCGACTC CTAAAGTCAA AAGTTGGCGG
CTCGCGTCTG CCCTGGGAGC GTCGCTCTGG AGTCGCTGAG GATTTCAGTT TTCAACCGCC

301 CGGGCGCCGG GCTCCCGCGG CTCTCACGG CCAGCTGCCTC CGTCGCCGC CGCAGCCAAG
GCCCGCGGCC CGAGGCGCGC GAGAGGTGCC GGCGACGGAG CGCAGCGGCC GCGTCGGTTC

+2 M Q F
Start met
=====
Kozak consensus
=====
Signal peptide
=====

361 GAGGGCAGGA GGGAGGGGGG TGGGGGCAGC GGAGGGAGGG GTGGGAAGCA CCATGCAGTT
CTCCCGTCCT CCCTCCCCC ACCCGCTCG CCTCCCTCCC CACCCCTCGT GGTACGTCAA

+2 V S W A T L L T L L V R D L A E M G S P
signal peptide cleavage site
=
Signal peptide Prodomain
=====

421 TGTATCCTGG GCCACACTGC TAACGCTCCT GGTGCGGGAC CTGGCCGAGA TGGGGAGCCC
ACATAGGACC CGGTGTGACG ATTGCGAGGA CCACGCCCTG GACCGGCTCT ACCCTCGGG

+2 D A A A A V R K D R L H P R Q V K L L E
Prodomain
=====

481 AGACGCCGCG GCGGCCGTGC GCAAGGACAG GCTGCACCCG AGGCAAGTGA AATTATTAGA
TCTGCGGCACG CGCCGGCACG CGTTCCGTG CGACGTGGC TCCGTTCACT TTAATAATCT

+2 T L S E Y E I V S P I R V N A L G E P F
Prodomain
=====

541 GACCCCTGAGC GAATACGAAA TCGTGTCTCC CATCGAGTG AACGCTCTCG GAGAACCTT
CTGGGACTCG CTTATGCTTT AGCACAGAGG GTAGGCTCAC TTGCGAGAGC CTCTGGAA

+2 P T N V H F K R T R R S I N S A T D P W
Prodomain
=====

601 TCCCCACGAAC GTCCACTTCA AAAGAACGCG ACGGAGCATT AACTCTGCCA CTGACCCCTG

+2 P A F A S S S S S S S T S S Q A H Y R L S
 Prodomain.
 ======
 661 GCCTGCCTTC GCCTCCTCCT CTTCCCTCCTC TACCTCCTCC CAGGCGCATT ACCGCCTCTC
 CGGACGGAAG CGGAGGAGGA GAAGGAGGAG ATGGAGGAGG GTCCGCGTAA TGGCGGAGAG

+2 A F G Q Q F L F N L T A N A G F I A P I
 Prodomain.
 ======
 721 TGCCCTCGGC CAGCAGTTTC TATTTAACCT CACCGCCAAT GCCGGATTTA TCGCTCCACT
 ACGGAAGCCG GTCGTCAAAG ATAAATTAGA GTGGCGGTTA CGGCCTAAAT AGCGAGGTGA

+2 F T V T L L G T P G V N Q T K F Y S E E
 Prodomain.
 ======
 781 GTTCACTGTC ACCCTCCTCG GGACGCCCGG GGTGAATCAG ACCAAGTTTT ATTCCGAAGA
 CAAGTGACAG TGGGAGGAGC CCTGCGGGCC CCACCTAGTC TGGTCAAAA TAAGGCTTCT

+2 E A E L K H C F Y K G Y V N T N S E H T
 Prodomain
 ======
 841 GGAAGCGGAA CTCAAGCACT GTTTCTACAA AGGCTATGTC AATACCAACT CCGAGCACAC
 CCTTCGCTT GAGTCGTGA CAAAGAIGTT TCCGATACAG TTATGGTTGA GGCTCGTGTG

+2 A V I S L C S G M L G T F R S H D G D Y
 Prodomain
 ======
 901 GGCGTCACTC AGCCTCTGCT CAGGAATGCT GGGCACATTG CGGTCTCATG ATGGGGATTA
 CCGGCAGTAG TCGGAGACGA GTCCTTACGA CCCGTGTAAG GCCAGAGTAC TACCCCTAAT

+2 F I E P L Q S M D E Q E D E E E Q N K P
 Prodomain
 ======
 961 TTTTATTGAA CCACTACAGT CTATGGATGA ACAAGAAGAT GAAGAGGAAC AAAACAAACC
 AAAATAACTT GGTGATGTCA GATACTTA TGTCTTCTA CTTCCTTGTG TTTTGTGTTGG

+2 H I I Y R R S A P Q R E P S T G R H A C
 Prodomain
 ======
 1021 CCACATCATT TATAGGCAGCA GCGCCCCCCC GAGAGAGCCC TCAACAGGAA GGCAATGCATG
 GGTGTAGTAA ATATCCCGGT CGCGGGGGGT CTCTCTCGGG AGTTGTCCTT CCGTACGTAC

+2 D T S E H K N R H S K D K K K T R A R K
 Prodomain
 ======
 1081 TGACACCTCA GAACACAAAA ATAGGCACAG TAAAGACAAG AAGAAAACCA GAGCAAGAAA
 ACTGTGGAGT CTTGTGTTT TATCCGTGTC ATTCTGTTC TTCTTTGGT CTCGTTCTT

+2 W G E R I N L A G D V A A L N S G L A T
 Prodomain
 ======
 1141 ATGGGGAGAA AGGATTAACC TGGCTGGTGA CGTAGCAGCA TAAACAGCG GCTTAGCAAC
 TACCCCTCTT TCCTAATTGG ACCGACCACT GCATCGTCGT AATTGTGCGC CGAATCGTTG

-2 E A F S A Y G N K T D N T R E K R T H F
Prodomain.

=====

1201 AGAGGCATTT TCTGCTTATG GTAATAAGAC GGACAAACACA AGAGAAAAGA GGACCCACAG
TCTCCGTAAA AGACGAATAC CATTATTCTG CCTGGTGTGT TCTCTTTCT CCTGGGTGTC

+2 R T K R F L S Y P R F V E V L V V A D N
Furin Cleavage site

=

Prodomain. Metalloproteinase domain.
=====

1261 AAGGACAAAA CGTTTTTAT CCTATCCACG GTTTGTAGAA GTCTGGTGG TGGCAGACAA
TTCCTGTTT GCAAAAATA GGATAGGTGC CAAACATCTT CAGAACCAACC ACCGTCTGT

+2 R M V S Y H G E N L Q H Y I L T L M S I
Metalloproteinase domain

=====

1321 CAGAATGGTT TCATACCATG GAGAAAACCT TCAACACTAT ATTTTAACCT TAATGTCAAT
GTCTTACCAA AGTATGGTAC CTCTTTGGA AGTTGTGATA AAAATTGAA ATTACAGTTA

+2 V A S I Y K D P S I G N L I N I V I V N
Metalloproteinase domain

=====

1381 TGTAGCCTCT ATCTATAAAAG ACCCAAGTAT TGGAAATTAA ATTAATATTG TTATTGTGAA
ACATCGGAGA TAGATATTTC TGGGTTCATAC ACCTTAAAT TAATTATAAC AATAACACTT

+2 L I V I H N E Q D G P S I S F N A Q T T
Metalloproteinase domain

=====

1441 CTTAATTGTG ATTCTATAATG AACAGGATGG GCCTTCCATA TCTTTTAATG CTCAGACAAAC
GAATTAACAC TAAGTATTAC TTGTCCTACC CGGAAGGTAT AGAAAATTAC GAGTCTGTG

+2 L K N F C Q W Q H S K N S P G G I H H D
Metalloproteinase domain

=====

1501 ATTAAAAAAAC TTTTGCCAGT GGCAGCATTC GAAGAACAGT CCAGGTGGAA TCCATCATGA
TAATTTTTG AAAACGGTCA CCGTCGTAAAG CTTCTGTCA GGTCCACCTT AGGTAGTACT

+2 T A V L L T R Q D I C R A H D K C D T I
Metalloproteinase domain

=====

1561 TACTGCTGTT CTCTAACAA GACAGGATAT CTGCAGAGCT CACGACAAAT GTGATAACCTT
ATGACGACAA GAGAATTGTT CTGTCCTATA GACGTCTGA GTGCTGTTA CACTATGGAA

+2 G L A E L G T I C D P Y R S C S I S E D
Metalloproteinase domain

=====

1621 AGGCCTGGCT GAACTGGAA CCATTTGTGA TCCCTATAGA AGCTGTTCTA TTAGTGAAGA
TCCGGACCGA CTTGACCCCTT GGTAAACACT AGGGATATCT TCGACAAGAT AATCACTCT

bioRxiv preprint doi: https://doi.org/10.1101/2023.09.11.553222; this version posted September 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

+2 S G L S T A F T I A H E L G H V F N M P
Zinc Binding Domain
=====

Metalloproteinase domain.
=====

1681 TAGTGGATTG AGTACAGCTT TTACGATCGC CCATGAGCTG GCCCATGTGT TAAACATGCC
ATCACCTAAC TCATGTCGAA AATGCTAGCG GGTACTCGAC CCGGTACACA AATTGTACGG

+2 H D D N N K C K E E G V K S P Q H V M A.
Zinc Binding Domain
=====

Metalloproteinase domain.
=====

1741 TCATGATGAC AACAAACAAAT GTAAAGAAGA AGGAGTTAAG AGTCCCCAGC ATGTCATGGC
AGTACTACTG TTGTTGTTA CATTCTTCT TCCTCAATTC TCAGGGTGC TACAGTACCG

+2 P T L N F Y T N P W M W S K C S R K Y I
Heparin Binding Motif
=====

Metalloproteinase domain. Heparin Binding Motif
=====

1801 TCCAAACACTG AACTTCTACA CCAACCCCTG GATGTGGTCA AAGTGTAGTC GAAAATATAT
AGGTTGTGAC TTGAAGATGT GGTTGGGAC CTACACCACT TTCACATCAG CTTTTATATA

+2 T E F L D T G Y G E C L L N E P E S R P
1861 CACTGAGTTT TTAGACACTG GTTATGGCGA GTGTTGGTAAACGAAACCTG AATCCAGACC
GTGACTCAAA AATCTGTGAC CAATACCGCT CACAAACGAA TTGCTTGGAC TTAGGTCTGG

+2 Y P L P V Q L P G I L Y N V N K Q C E L
Disintegrin domain
=====

1921 CTACCCCTTG CCTGTCCAAC TGCCAGGCAT CCTTACAAC GTGAATAAAC AATGTGAATT
GATGGGAAAC GGACAGGTTG ACGGTCCGTA GGAAATGTIG CACTTATTG TTACACTAA

+2 I F G P G S Q V C P Y M M Q C R R L W C
Disintegrin domain
=====

1981 GATTTTGGA CCAGGTTCTC AGGTGTGCC ATATATGATG CAGTGCAGAC GGCTCTGGTG
CTAAAAACCT GGTCCAAGAG TCCACACGGG TATATACTAC GTCACGTCTG CCGAGACCAC

+2 N N V N G V H K G C R T Q H T P W A D G
Disintegrin domain
=====

2041 CAATAACGTC AATGGAGTAC ACAAAAGGCTG CCGGACTCAG CACACACCTT GGGCCGATGG
GTTATTGCAG TTACCTCATG TGTTCCGAC GGCTGAGTC GTGTGTGGGA CCCGGCTACC

+2 T E C E P G K H C K Y G F C V P K E M D
Disintegrin domain
=====

2101 GACGGAGTGC GAGCCTGGAA AGCACTGCAA GTATGGATT TGTGTTCCA AAGAAATGGA
CTGCCTCACG CTCGGACCTT TCGTGACGTT CATACCTAAA ACACAAGGGT TTCTTACCT

. . . V P V T D G S W G S W S F F G T C S R ?
 Disintegrin domain TSPI
 ======
 2161 TGTCCCCGTG ACAGATGGT CCTGGGAAG TTGGAGTCCC TTGGAACCT GCTCCAGAAC
 ACAGGGGCAC TGTCTACCTA GGACCCCTTC AACCTCAGGG AACCTTGGG CGAGGTCTTG

 +2 C G G G I K T A I R E C N R P E P K N G
 TSPI
 ======
 2221 ATGTGGAGGG GGCATCAAAA CAGCCATTG AGAGTGCAC AGACCAGAAC CAAAAAAATGG
 TACACCTCCC CCGTAGTTT GTCGGTAAGC TCTCACGTTG TCTGGTCTTG GTTTTTTACC

 +2 G K Y C V G R R M K F K S C N T E P C L
 TSPI Spacer Region
 ======
 2281 TGGAAAATAC TGTGTAGGAC GTAGAATGAA ATTTAAGTCC TGCAACACGG AGCCATGTCT
 ACCTTTATG ACACATCCTG CATCTACTT TAAATTCAAGG ACGTITGTGCC TCGGTACAGA

 +2 K Q K R D F R D E Q C A H F D G K H F N
 Spacer Region
 ======
 2341 CAAGCAGAAG CGAGACTTC GAGATGAACA GTGTGCTCAC TTGACGGGA AGCATTTAA
 GTTCGTCTTC GCTCTGAAGG CTCTACTTGT CACACGAGTG AACTGCCCT TCGTAAAATT

 +2 I N G L L P N V R W V P K Y S G I L M K
 Spacer Region
 ======
 2401 CATCAACGGT CTGCTCCCCA ATGTGCGCTG GGTCCTAAA TACAGTGGAA TTCTGATGAA
 GTAGTTGCCA GACGAAGGGT TACACGCGAC CCAGGGATT ATGTCACCTT AAGACTACTT

 +2 D R C K L F C R V A G N T A Y Y Q L R D
 Spacer Region
 ======
 2461 GGACCGGTGC AAGTTGTTCT GCAGAGTGGC AGGAACACA GCCTACTATC AGCTTCGAGA
 CCTGGCCACG TTCAACAAGA CGTCTCACCG TCCCTGTGT CGGATGATAG TCGAAGCTCT

 +2 R V I D G T P C G Q D T N D I C V Q G L
 Spacer Region
 ======
 2521 CAGAGTGATA GATGGAACTC CTTGTGCCA GGACACAAAT GATATCTGTG TCCAGGGCCT
 GTCTCACTAT CTACCTTGAG GAACACCGGT CCTGTGTTA CTATAGACAC AGGTCCCGGA

 +2 C R Q A G C D H V L N S K A R R D K C G
 Spacer Region
 ======
 2581 TTGCCGGCAA GCTGGATGCG ATCATGTTT AAACTAAAAA GCCCGGAGAG ATAAATGTGG
 AACGGCCGTT CGACCTACGC TAGTACAAAA TTGAGTTT CGGGCCTCTC TATTTACACC

 +2 V C G G D N S S C K T V A G T F N T V H
 Spacer Region
 ======
 2641 GGTGGTGGT GGCATAATT CTTCATGCAA AACAGTGGCA GGAACATTAA ATACAGTACA
 CCAAACACCA CCGCTATTAA GAAGTACGTT TTGTCACCGT CCTTGAAAT TATGTCATGT

+2 Y G Y N T V V R I P A G A T N I D V R C
 Spacer Region

=====

01 TTATGGTTAC AATACTGTGG TCCGAATTCC AGCTGGTGC ACCAATATTG ATGTGCCGA
 AATAACCAATG TTATGACACC AGGCTTAAGG TCGACCACGA TGTTTATAAC TACACGCCGT

+2 H S F S G E T D D D N Y L A L S S S K G
 Spacer Region

=====

51 GCACAGTTTC TCAGGGGAAA CAGACGATGA CAACTACTTA GCTTATCAA GCAGTAAAGG
 CGTGTCAAAG AGTCCCCTTT GTCTGCTACT GTTGATGAAT CGAAATAGTT CGTCATTTC

+2 E F L L N G N F V V T M A K R E I R I G
 Spacer Region

=====

21 TGAATTCTTG CAAATGGAA ACTTTGTTGT CACAATGGCC AAAAGGGAAA TTCGCATTGG
 ACTTAAGAAC GATTTACCTT TGAAACAAACA GTGTTACCGG TTTTCCCTTT AAGCGTAACC

+2 N A V V E Y S G S E T A V E R I N S T D
 Spacer Region

=====

31 GAATGCTGTG GTAGAGTACA GTGGGTCCGA GACTGCCGT AAAAAAGGTTA ACTCAACAGA
 CTTACGACAC CATCTCATGT CACCCAGGCT CTGACGGCAT CTTCTTAAT TGAGTTGCTT

+2 R I E Q E L L L Q V L S V G K L Y N P D
 Spacer Region

=====

41 TCGCATTGAG CAAGAACCTT TGCTTCAGGT TTTGTCGGTG GAAAAGTTGT ACAACCCCCGA
 AGCGTAACTC GTTCTTGAAA ACGAAGTCCA AAACAGCCAC CCTTCAACA TGTTGGGGCT

+2 V R Y S F N I P I E D K P Q Q F Y W N S
 Spacer Region

=====

51 TGTACGCTAT TCTTCAATA TTCCAATG AAGATAAACCT CAGCAGTTT ACTGGAACAG
 ACATGCGATA AGAAAGTTAT AAGGTTAATCT TCTATTGGA GTCGTCAAAA TGACCTTGTC

+2 H G P W Q A C S K P C Q G E R K R K L V
 Spacer Region

=====

61 TCATGGGCCA TGGCAAGCAT GCAGTAACCC CTGCCAAGGG GAACGGAAAC GAAAACCTGT
 AGTACCCGGT ACCGTTCGTA CGTCATTGG GACGGTCCCC CTTGCCCTTG CTTTGACAC

+2 C T R E S D Q L T V S D Q R C D R L F C
 Spacer Region

=====

21 TTGCACCAGG GAATCTGATC AGCTTACTGT TTCTGATCAA AGATGCGATC GGCTGCCCA
 AACGTGGTCC CTTAGACTAG TCGAATGACA AAGACTAGTT TCTACGCTAG CCGACGGGGT

+2 P G H I T E P C G T D C D L R W H V A
 Spacer Region

=====

81 GCCTGGACAC ATTACTGAAC CCTGTGGTAC AGACTGTGAC CTGAGGTGGC ATGTTGCCA
 CGGACCTGTG TAATGACTTG GGACACCATG TCTGACACTG GACTCCACCG TACAACCGT

Y. H. Li et al. / Journal of Molecular Biology 362 (2006) 102–114

+2 R S E C S A Q C G L G Y R T L D I Y C A
Spacer Region

3241 CAGGAGTGAA TGTAGTGCC AGTGTGGCTT GGGTTACCGC ACATTGGACA TCTACTGTGC
GTCCCTCACTT ACATCACGGG TCACACCGAA CCCAATGGCG TGTAACCTGT AGATGACACG

+2 K Y S R L D G K T E K V D D G F C S S H
Spacer Region

3301 CAAATATAGC AGGCTGGATG GGAAGACTGA GAAGGTTGAT GATGGTTTT GCAGCAGCCA
GTTTATATCG TCCGACCTAC CCTTCTGACT CTTCCAACTA CTACCAAAAA CGTCGTCGGT

+2 P K F S N R E K C S G E C N T G G W R Y
Spacer Region

3361 TCCAAACCA AGCAACCGTG AAAATGCTC AGGGGAATGT AACACGGGTG GCTGGCGCTA
AGGGTTTGGT TCGTTGGCAC TTTTACGAG TCCCCCTACA TTGTGCCAC CGACCGCGAT

+2 S A W T E C S K S C D G G T Q R R R A I
Spacer Region

TSP2

3421 TTCTGCCTGG ACTGAATGTT CAAAAAGCTG TGACCGGTGG ACCCAGAGGA GAAGGGCTAT
AACACGGACC TGACTTACAA GTTTTCGAC ACTGCCACCC TGGGTCTCCT CTTCCCGATA

+2 C V N T R N D V L D D S K C T H Q E K V
TSP2

3481 TTGTGTCAAT ACCCGAAATG ATGTACTGGA TGACAGCAA TGCACACATC AAGAGAAAAGT
AACACAGTTA TGGGCTTTAC TACATGACCT ACTGTCGTTT ACGTGTGTAG TTCTCTTCA

+2 T I Q R C S E F P C P Q W K S G D W S E
TSP2

TSP3

3541 TACCATTCAG AGGTGCAGTG AGTCCCTTG TCCACAGTGG AAATCTGGAG ACTGGTCAGA
ATGGTAAGTC TCCACGTAC TCAAGGGAAC AGGTGTCAAC TTTAGACCTC TGACCAAGTCT

+2 C L V T C G K G H K H R Q V W C Q F G E
TSP3

TSP3

3601 GTGCTGGTC ACCTGTGGAA AAGGGATAA GCACCGCCAG GTCTGGTGT AGTTGGTGA
CACGAACCGAG TGGACACCTT TCCCCGTATT CGTGGCGGTG CAGACCACAG TCAAACCACT

+2 D R L N D R M C D F E T K P T S M Q T C
TSP3

TSP3

3661 AGATCGATTA AATGATAGAA TGTGTGACCC TGAGACCAAG CCAACATCTA TGCAGACTTG
TCTAGCTAAT TTACTATCTT ACACACTGGG ACTCTGGTTC GGTTGTAGAT ACGTCTGAAC

+2 Q Q P E C A S W Q A G P W G Q C S V T C
TSP3

TSP4

3721 TCAGCAGCCG GAATGTGCAT CCTGGCAGGC GGGTCCCTGG GGACAGTGCA GTGTCACTTG
AGTCGTCGGC CTTACACGTA GGACCGTCCG CCCAGGGACC CCTGTCACGT CACAGTGAAC

PAPUA NEW GUINEA

+1 G Q G Y Q L R A V K C I I G T Y M S V V
TSP4

=====
3781 TGGACAGGGA TACCAAGCTAA GAGCAGTGAA ATGCATCATT GGGACTTATA TGTCAGTGGT
ACCTGTCCCT ATGGTCGATT CTCGTCACTT TACGTAGTAA CCCTGAATAT ACAGTCACCA

+2 D D N D C N A A T R P T D T Q D C E L F
TSP4

=====
3841 AGATGACAAT GACTGTAATG CAGCAACTAG ACCAACTGAT ACCCAGGACT GTGAATTACC
TCTACTGTTA CTGACATTAC GTCGTTGATC TGGTTGACTA TGGTCCTGA CACTTAATGG

+2 S C H P P P A A P E T R R S T Y S A P R
TSP4

=====

3901 ATCATGTCAT CCTCCCCAG CTGCCCGGA AACGAGGAGA AGCACATACA GTGCACCAAG
TAGTACAGTA GGAGGGGGTC GACGGGGCCT TTGCTCCTCT TCGTGTATGT CACGTGGTC

+2 T Q W R F G S W T P C S A T C G K G T R
TSP5

=====
3961 AACCCAGTGG CGATTGGGT CTGGACCCCC ATGCTCAGCC ACTTGTGGGA AAGGTACCCG
TTGGGTACCC GCTAAACCCA GAACCTGGGG TAGCAGTCGG TGAACACCCCT TTCCATGGGC

+2 M R Y V S C R D E N G S V A D E S A C A
TSP5

=====
4021 GATGAGATAAC GTCAGCTGCC GAGATGAGAA TGGCTCTGTG GCTGACGAGA GTGCCTGTGC
CTACTCTATG CAGTCGACGG CTCTACTCTT ACCGAGACAC CGACTGCTCT CACGGACACG

+2 T L P R P V A K E E C S V T P C G Q W K
TSP5 TSP6

=====
4081 TACCTGCCT AGACCAGTGG CAAAGGAAGA ATGTTCTGTG ACACCCCTGTG GGCAATGGAA
ATGGGACGGA TCTGGTCACC GTTTCCTCT TACAAGACAC TGTGGGACAC CCGTTACCTT

+2 A L D W S S C S V T C G Q G R A T R Q V
TSP6

=====
4141 GGCTTGGAC TGGAGCTCT GCTCTGTGAC CTGTGGCAA GGTAGGGCAA CCCGGCAAGT
CCGGAACCTG ACCTCGAGAA CGAGACACTG GACACCCGTT CCATCCCGTT GGGCCGTTCA

+2 M C V N Y S D H V I D R S E C D Q D Y I
TSP6

=====
4201 GATGTGTGTC AACTACAGT ACCACGTGAT CGATCGGAGT GAGTGTGACCC AGGATTATAT
CTACACACAG TTGATGTAC TGGTGCACCA GCTAGCCTCA CTCACACTGG TCCTAATATA

+2 P E T D Q D C S M S P C P Q R T P D S G
TSP6

=====
4261 CCCAGAAACT GACCAGGACT GTTCCATGTC ACCATGCCCT CAAAGGACCC CAGACAGTGG
GGGTCTTGA CTGGTCCTGA CAAGGTACAG TGGTACGGGA GTTCTGGGG GTCTGTACCA

+2 L A Q H P F Q N E D Y R F R S A S P S F
 4321 CTTAGCTCAG CACCCCTTCC AAAATGAGGA CTATCGTCCC CGGAGCGCCA GCCCCAGCGC
 GAATCGAGTC GTGGGGAAAGG TTTTACTCCT GATAGCAGGG GCCTCGCGGT CGGGGTCGGC

 +2 T H V L G G N Q W R T G P W G A C S S T
 TSP7

 =====
 4381 CACCCATGTG CTCGGTGCAA ACCAGTGGAG AACTGGCCCC TGGGGAGCAT GTTCCAGTAC
 GTGGGTACAC GAGCCACCTT TGGTCACCTC TTGACCGGGG ACCCCTCGTA CAAGGTCAATG

 +2 C A G G S Q R R V V V C Q D E N G Y T A
 TSP7

 =====
 4441 CTGTGCTGGC GGATCCCAGC GGCGTGTGT TGTATGTCAG GATGAAAATG GATACACCGC
 GACACGACCG CCTAGGGTCG CGGCACAACA ACATACAGTC CTACTTTAC CTATGTGGCG

 +2 N D C V E R I K P D E Q R A C E S G P C
 TSP7

 =====
 4501 AACAGACTGT GTGGAGAGAA TAAAAACCTGA TGAGCAAAGA GCCTGTGAAT CCGGCCCTTG
 TTTGCTGACA CACCTCTCTT ATTTGGACT ACTCGTTTCT CGGACACTTA GGCGGGAAAC

 +2 P Q W A Y G N W G E C T K L C G G G I R
 TSP7

 =====
 4561 TCCTCAGTGG GCTTATGGCA ACTGGGGAGA GTGCACTAAG CTGTGTGGTG GAGGCATAAG
 AGGAGTCACC CGAACACCGT TGACCCCTCT CACGTGATTG GACACACCACT CTCGTATTG

 +2 T R L V V C Q R S N G E R F P D L S C E
 TSP8

 =====
 4621 AACAAAGACTG GTGGTCTGTC AGCGGTCCAA CGGTGAACGG TTTCCAGATT TGAGCTGTGA
 TTGTTCTGAC CACCAGACAG TCGCCAGGTT GCACATTGCC AAAGGTCTAA ACTCGACACT

 +2 I L D K P P D R E Q C N T H A C P H D A
 TSP8

 =====
 4681 AATTCTTGAT AACCTCCCG ATCGTGAGCA GTGTAACACA CATGCTTGTC CACACGACGC
 TTAAGAACTA TTGGGAGGGC TAGCACTCGT CACATTGTGT GTACGAACAG GTGTGCTGCG

 +2 A W S T G P W S S C S V S C G R G H K Q
 TSP9

 =====
 4741 TGCATGGAGT ACTGGCCCTT GGAGCTCGTG TTCTGTCTCT TGTGGTCGAG GGCATAAAACA
 ACGTACCTCA TGACCGGGAA CCTCGAGCAC AAGACAGAGA ACACCACTCT CCGTATTGTT

 +2 R N V Y C M A K D G S H L E S D Y C K H
 TSP9

 =====
 4801 ACGAAATGTT TACTGCATGG CAAAAGATGG AAGCCATTAA GAAAGTGATT ACTGTAAGCA
 TGTTTACAA ATGACGTACC GTTTCTACC TTCGGTAAAT CTTTCACTAA TGACATTCTG

 +2 L A K P H G H R K C R G G R C P K W K A
 TSP10

4861 CCTGGCTAAG CCACATGGGC ACAGAAAGTG CGGAGGAGGA AGATGCCCA AATGGAAAGC
 GGACCGATTC GGTGTACCCG TGTCTTICAC GGCTCCTCCT TCTACGGGGT TTACCTTCG
 +2 G A W S Q C S V S C G R G V Q Q R H V G
 TSP10
 =====
 4921 TGGCGCTTGG AGTCAGTGCT CTGTGTCTG TGCCCGAGGC GTACAGCAGA GGCATGTGGG
 ACCCGAACC TCAGTCACGA GACACAGGAC ACCGGCTCCG CATGTCGTCT CGGTACACCC
 +2 C Q I G T H K I A R E T E C N P Y T R F
 TSP10
 =====
 4981 CTGTCAGATC GGAACACACA AAATAGCCAG AGAGACCGAG TGCAACCAT ACACCAAGACC
 GACAGTCTAG CCTTGTGTGT TTTATCGGTC TCTCTGGCTC ACGTTGGGTA TGTGGTCTGG
 +2 E S E R D C Q G P R C P L Y T W R A E E
 TSP11
 =====
 5041 GGAGTCGGAA CGCGACTGCC AAGGCCACAG GTGTCCCCTC TACACTTGGGA GGGCAGAGGA
 CCTCAGCCTT GCGCTGACGG TTCCGGGTGC CACAGGGGAG ATGTGAACCT CCCGTCCTCCT
 +2 W Q E C T K T C G E G S R Y R K V V C V
 TSP11
 =====
 5101 ATGGCAAGAA TGCACCAAGA CCTGCGGCGA AGGCTCCAGG TACCGCAAGG TGGTGTGTGT
 TACCGTTCTT ACGTGGTTCT GGACGCCGCT TCCGAGGTCC ATGGCGTTCC ACCACACACA
 +2 D D N K N E V H G A R C D V S K R P V D
 TSP11
 =====
 5161 GGATGACAAC AAAAACGAGG TGCGATGGGC ACGCTGTGAC GTGAGCAAGC GGCGGGTGGG
 CCTACTGTIG TTTTGCTCC ACGTACCCCG TCGCACACTG CACTCGTTCG CGGGCACACT
 +2 R E S C S L Q P C E Y V W I T G E W S E
 TSP11
 =====
 5221 CCCGTAAAGC TGTAGTTGCA AACCTTGGCA GTATGTCTGG ATCACAGGAG AATGGTCAGA
 GGCACCTTCG ACATCAAACG TTGGGACGCT CATAACAGACC TAGTGTCTC TTACCAAGTCT
 +2 V P S W E L *
 * Termination codon
 ====
 5281 GGTACCGTCC TGGGAACGT AACCATCGTC AGCTCAGCCA TGGCCTGAGA GTGGCAGAGG
 CCATGGCAGG ACCCTTGACA TTGGTAGCAG TCGAGTCGGT ACCGGACTCT CACCGTCTCC
 5341 GATGAGTGGG GGGATGAGTG CAGGAATGTG GGAGACTTGA GGCTACCCGC CCGATTTGCC
 CTACTCACCT CCCTACTCAC GTCTTACAC CCTCTGAACCT CCGATGGCG GGCTAAACCG
 5401 ACTGTGAAC TGTGTCTTGC TGACAAAGTCC TCAGCTTCC CAAGCTAGAA TTCCCTGTAT
 TGACACTTGA CACACAAAAG ACTGTTCAAG AGTCGAAAGG GTTCGATCTT AAGGAACATA
 5461 GCAAAGCGGG AGAGATGTAA GAGATGGTCT CTAAGTCCCT TCAGGTCTAC ATTCTGTGAT
 CGTTTCGCC TCTCTACATT CTCTACAGA GATTCAAGGG AAGTCCAGATG TAAGACACTA

5521 TCACCTTGAT GTCCTATTGG CATAAAGAAG AAATTATTAC ACGGGGCTGCA AACTCATAGC
AGTGGAACTA CAGGATAACC GTATTTCTTC TTTAATAATG TCCCCGACGT TTGAGTATCC

5581 ATGCTGTGAG GTGCCTGAAG ACAGTTAAGT ATAAGAAAAT ATTGTAGTGC CAGGGATACA
TACGACACTC CACGGACTTC TGTCATTCA TATTCTTTA TAACATCACG GTCCCTATGT

5641 ACAAGGAGAG ATGGCAACTG TGACAAACTA GCACATGCTG TGTGAAGGGA GCAGAACTC
TGTTCCTCTC TACCGTTGAC ACTGTTGAT CGTGTACGAC ACACCTCCCT CGTCTTAGAG

5701 TTTCACTCCA GCTGTGGCCA TGCAGAAATG TGGTCTAGCG TTACCAAGACC TGATTTTCA
AAAGTGAGGT CGACACCGGT ACGTCTTAC ACCAGATCGC AATGGTCTGG ACTAAAAAGT

5761 AGAGAGGCTA AAAATCTGGA CTAGTATGTG AGATTTCTA ACTTGAAAAT GGGGGCTGAA
TCTCTCCGAT TTTTAGACCT GATCATAACAC TCTAAAGGAT TGAACCTTTA CCCCCGACTT

5821 ATTTTGTTT TAAACATT GTAAGGGCA AACAAACCCC TTCATGAAC CAGATGTGTT
TAAACCAA AATTTGTAA CATTCCCCGT TTGTTGGGG AAAGTACTTG GTCTACACAA

5881 GTGCCTGTT AACAAACAGC TTCAGAGGAA GAAAATAATT TTCTATAATA TCCGAAAGTAT
CACGGACAAA TTGTTTGTG AAGTCTCCTT CTTTATTAA AAGATATTAT AGGCTTCATA

5941 CTCAAGTACC ATTTTTCAT ATATCTCCT GTGCACAATG CTTATCTAGA CCCTTTTAA
GAGTTCAIGG TAAAAAAAGTA TATAGAAGGA CACGTGTAC GAATAGATCT GGGAAAATT

PolyA Site

=====

6001 TGTAATAAA CCAGTAGTAA TCAT
ACCATTATTT GGTCAATCATT AGTA

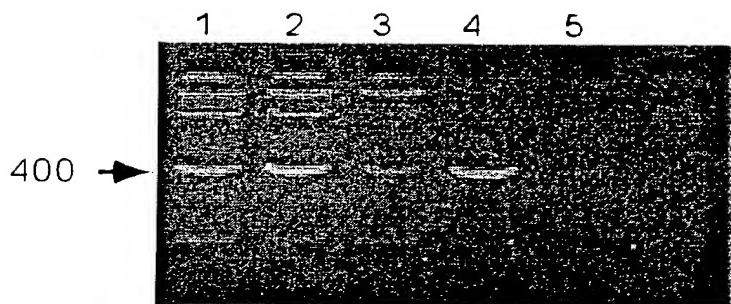
Fig. 4

Metalloproteinase Domains Alignment of TS-SI v. ADAMTS Family

SCANNED# 24

		Percent Homology (SI vs)
30		
hADAMTS-4 (AB014588)	(212) REASLSREVETLLVADDUKMAAFHS-AGLERYLLAANAAAAKHPPEIRNPNSLIVYTRLV	360
hADAMTS-5 (AF142099)	(261) RSISRARQVELLVAADSMARL--R-GLOHMDLALLSIANRISHA-ENHEDRAAVKVV	46
hADAMTS-1 (AF060152)	(235) REVSSHRREVETLLVADSMARL--SGKMEEL--STARYKHPPEIRNPNSLIVYTRLV	44
hADAMTS-8 (AF060153)	(214) REVSEREVETLLVADSMARL--ADIONHEEITPSVARYKHPPEIRNPNSLIVYTRLV	50
hADAMTS-2 (AJ003125)	(260) HAADDYDNEEVLGDDIVOKEKEHDKXKXLELNNVNEEDDESGAHINWVIVVII	49
nADAMTS-SI	(287) RELSREVETLLVADNNSVSHS-ENQQLXDEPNTIVASAKDFCQNLNQVNLNLI	33
Consensus	(301) RF S REVETLLVAD SMA FHK GLOXYLTIASIARYKHPSI N INLVVK	420
hADAMTS-4 (AB014588)	(271) DFGSGREGBQGP-SAAQVRSFATRGGINTBEDDPHNTTILFNSODCGGS-TGD	361
hADAMTS-5 (AF142099)	(320) VIGDKDKSLEWK-NATIENKQVQHONQGDIEERVAVLIFREEDGCHH-SED	420
hADAMTS-1 (AF060152)	(294) VHDQXQGETPS-NATIENKQVQHONQGDIEERVAVLIFREEDGCHH-TGD	420
hADAMTS-8 (AF060153)	(273) ISEDERKGRVEYD-NGS-AHNRERWVORRETOSDRDPEHYTTEVTFQFGELEQ	420
hADAMTS-2 (AJ003125)	(320) IISYGRSMSLIEGNPSOENVRYAYLOCKEDGIDGDEKTHAATFQFQPGS---	420
nADAMTS-SI	(346) VHNEDQPSHSF-NATIENKQVQHONQGDIEERVAVLIFREEDGCHH-TGD	420
Consensus	(361) VL DE GPEVS GNAA TLRNFC WQ N P D H EHYTAILTQDLGC G CD	420
hADAMTS-4 (AB014588)	(329) TISMADGTVIADARSSEEDDEQDQDPAHLLVYFNMLINN-SKATISLNGPLSTS	480
hADAMTS-5 (AF142099)	(378) TDEMADGTVIADARSSEEDDEQDQDPAHLLVYFNMLINN-SKATISLNGPLSTS	480
hADAMTS-1 (AF060152)	(352) TLEMADGTVIADRSSESDDEQDQDPAHLLVYFNMLINN-SKATISLNGPLSTS	480
hADAMTS-8 (AF060153)	(332) TLEVADGTVIADINKSOSYTEDDEQDQDPAHLLVYFNMLINN-SKATISLNGPLSTS	480
hADAMTS-2 (AJ003125)	(376) MQYAPTCGHVRSQPHNDLTSVVAWAEVTFVGMENQGNGNGDEVRLGS--	480
nADAMTS-SI	(402) TIGMADGTVIADYRSEEDDQDQDPAHLLVYFNMLINN-SKATISLNGPLSTS	480
Consensus	(421) TLGMAADGTVIADDSVQEDDQDQDPAHLLVYFNMLINN-SKATISLNGPLSTS	480
hADAMTS-4 (AB014588)	(388) RHYMAPYIAHVDPPEPEPSARPDLE DUGICHTLQKKA PLHRYWTFP--KDD	481
hADAMTS-5 (AF142099)	(437) -RHYMAPSITSIDASKPKSKTSATWEEFDQDHGNCNLQDPLRK QIGGEELPS--QTD	481
hADAMTS-1 (AF060152)	(411) -HMMASQISLNLDISQPSRASYMUSFEDNGEENDKQN-BIOLBQDGQS--TSID	481
hADAMTS-8 (AF060153)	(391) -HYMAPFQVHLNQTLPSPSAMYIPLADGQHGDDQDIAAGA ALBLEPQGPERALAQ	481
hADAMTS-2 (AJ003125)	(434) -DMAPYQOAFHRFHRSQOEEHHS-YDQDIDDEFADWALQPE--LHS	481
nADAMTS-SI	(460) -HYMAPFNFYTPWMSKRSKYUEEEDSSEGEENEPSEHYRPHQES--ILN	481
Consensus	(481) RHVMAP L D PNSPCSA ITEFLD GHG CLID P ARP PLP LPGRM YD	541

FIG 5



© 2000 by Marcel Dekker, Inc. All rights reserved.

Figure 6

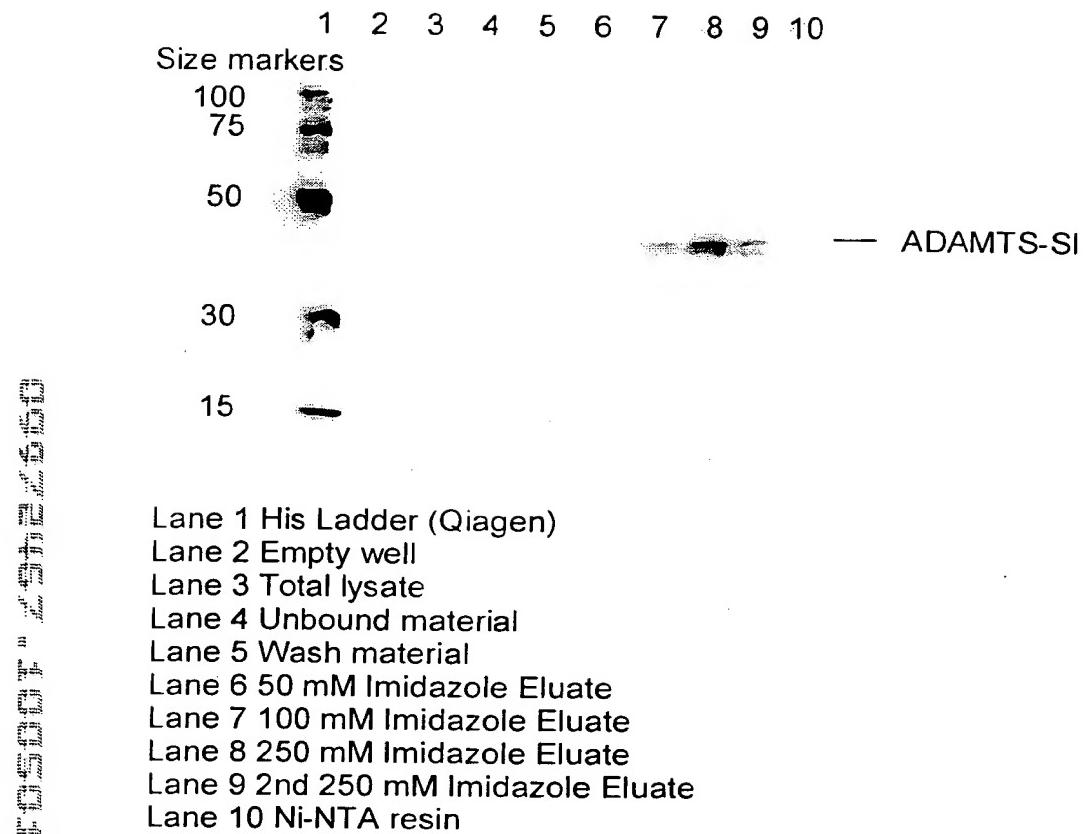


FIG. 7

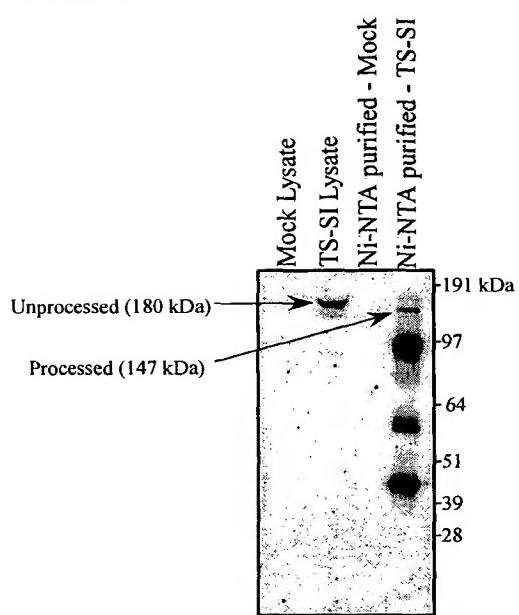
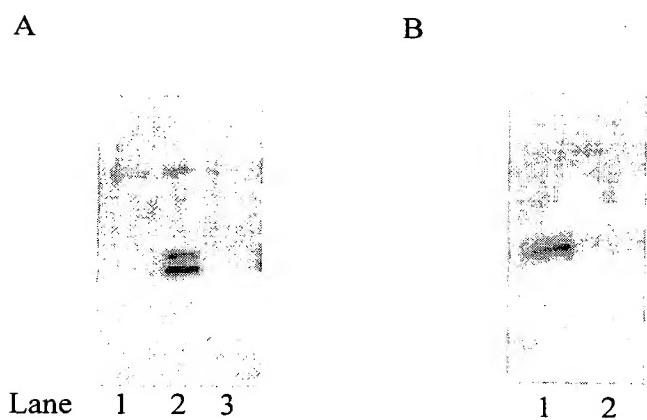


FIG. 8



Western Blot analysis of ADAMTS-4 and ADAMTS-SI cleavage of aggrecanase substrate. A, lane 1 = uncleaved substrate, lane 2 = ADAMTS-4, lane 3 = ADAMTS-4 + 1 μ M compound I. B, lane 1 = ADAMTS-SI and lane 2 = ADAMTS-SI + 1 μ M compound I.